4   or all inputs are missing or are otherwise unknown, the method [steps] comprising:

          (1)   presenting a collection of training data comprising examples of input

6   values that are available to the model together with [the] corresponding

7   desired output value(s) that the model is intended to predict;

8   [and]

9   (2)   generating a plurality of subordinate models, that together comprise an

10   overall model, in such a way that:

11   a)   each subordinate model has an associated set of application

12   conditions that must be satisfied in order to apply the

13   subordinate model when making predictions, the application

14   conditions comprising:

15   i)   tests for missing values for all, some, or none

16   of the inputs,

17   and

18   ii)   tests on the values of all, some, or none of the

19   inputs that are applicable when the values of

20   the inputs mentioned in the tests have known

21   values;

22   and

23   b)   for at least one subordinate model, the training cases used in

24   the construction of that subordinate model include some

25   cases that indirectly satisfy the application conditions [in the

26   sense] such that the application conditions are satisfied only

27   after replacing one or more known data values in these

28   training cases with missing values; and

29   3) outputting a specification of at least one of said subordinate models thus generated

30   and making a prediction based on said at least one of said subordinate models thus-generated.

2.   (Amended) A device according to claim 1, wherein step (2) comprises

generating a plurality of subordinate models [with the further requirement]

such that the plurality [CANNOT] cannot be arranged into a decision-tree hierarchy in such a way that:

(1)   each branch of the tree corresponds to a test on the values of one or more data fields that can be satisfied only when those data fields have known values;

(2)   each leaf of the tree corresponds to a subordinate model whose application conditions are defined by the conjunction of the tests along the branches that lead from the root node of the tree to the leaf node;

(3)   the root node of the tree corresponds to a subordinate model whose application conditions [consist of] include missing-value tests for the data fields mentioned in the tests associated with the tree branches that emanate from the root node;
         and

(4)   each interior node of the tree other than the root node corresponds to a subordinate model whose application conditions are defined by the conjunction of the tests along the branches that lead from the root node of the tree to the interior node, together with missing-value tests for the data fields mentioned in the tests associated with the tree branches that emanate from the interior node.

(Please add the following new claims:)

–3. The program storage device according to claim 1, wherein, when an additional data field is incorporated into the construction of a subordinate model, an alternate subordinate model is constructed for use when said additional data field has a missing value.

4. The program storage device according to claim 1, wherein a missing value is estimated by performing a prediction based on the known data values.

5. The program storage device according to claim 1, wherein each subordinate model has an

application condition that must be satisfied for said each subordinate model to be applied, and

wherein said application condition includes at least one of the values to be input to the model being missing.

6. The program storage device according to claim 1, wherein said outputting comprises outputting a specification of a plurality of subordinate models and their associated application conditions, and reading said specification being readable by the machine.

7. The program storage device according to claim 1, wherein said values are missing at random.

8. The program storage device according to claim 1, wherein based on said data collection, it is determined whether missing data values are missing at random or whether missing values convey information.

9. The program storage device according to claim 1, wherein a determination of randomness of missing values is made by examining the data values present.

10. The program storage device according to claim 1, wherein statistical tests are employed to determine randomness of missing values.

11. The program storage device according to claim 1, wherein randomness of missing values is assessed with a cross-validation technique.

12. The program storage device according to claim 11, wherein applying the cross validation technique comprises:

selecting and holding aside portions of the training cases that directly satisfy the application conditions of a subordinate model for validation purposes;

constructing first and second models using remaining training cases that directly satisfy the application conditions but were not held aside, such that one of the first and second

models is constructed based only on the remaining cases and the second model is constructed based on the remaining cases plus the training cases that indirectly satisfy the application conditions;

estimating prediction errors of the first and second models by applying the models to the training cases held aside for validation purposes;

if a predictive accuracy of the first model is greater than that of the second model with a predetermined sufficiently high statistical significance, then assuming that missing values in the relevant fields are informative and the subordinate model should be constructed only from those training cases that directly satisfy the application conditions of the subordinate model; and

if a predictive accuracy of the first model is greater than that of the second model with a predetermined sufficiently high statistical significance, then missing values are treated as random events and the training cases that directly or indirectly satisfy the application conditions are used in the construction of the subordinate model.

13. The program storage device according to claim 12, wherein the cross-validation method further comprises:

if a subordinate model is constructed for use when two or more data fields have missing values, then missing values of some of these data fields are treated as missing at random and others of said data fields are treated as informative,

wherein the training cases constructing the subordinate model includes those that directly satisfy the application conditions of the subordinate model together with those that indirectly satisfy the application conditions when known data values are replaced with missing values, but only for those data fields for which missing values are to be treated as missing at random.

14. The program storage device according to claim 13, wherein determining whether said missing values should be treated as missing at random or which should be treated as informative, includes:

constructing a model assuming that all missing values are to be treated as informative,

such that the model is constructed from those training cases that directly satisfy the application conditions of the subordinate model but are not being held aside for validation purposes, said model being termed the "current model";

for each missing value in the "current model" that is treated as informative, constructing another model that treats that missing value as missing at random while treating all other missing values in the same manner as the "current model";

of the new models, choosing the one model that yields the greatest predictive accuracy on the training cases defined in said constructing that were used to construct the first "current model," and calling this new model the "current model";

repeating the constructing of the another model and the choosing until all missing values are treated as missing at random by the "current model";

of all "current models" obtained in the constructing of the "current model" and choosing, choosing the model that yields the greatest predictive accuracy on the training cases held aside for validation purposes, and calling this model the "best model"; and

constructing the subordinate model, without holding training cases aside for validation purposes, using the same treatments of missing values used in the construction of the "best model."

15.   The program storage device according to claim 1, wherein a determination as to how to treat missing values for subordinate models is deferred.

16.   The program storage device according to claim 1, wherein if a top-down method is employed to construct the subordinate models, then the plurality of models include a single subordinate model that does not use any data fields as input and which has an application condition that is always true.

17.   The program storage device according to claim 1, wherein if a bottom-up method is employed to construct the subordinate model, then the plurality of models include a plurality of subordinate models and application conditions, the application conditions covering all possible combinations of values of the data fields.